

Examining the Correlation between the Types of Businesses near BLUEbikes Stations and the Proportion of Subscribers using those Stations

Makana Burch, '25, Data Science Major Capstone

Question

Is there a relationship between the proportion of subscribers ending their trips at a station and the types of businesses within a half-mile radius of that station (office, food, recreation, etc.) in 2019? Do users in different age groups end their rides at stations with significantly different types of businesses nearby?

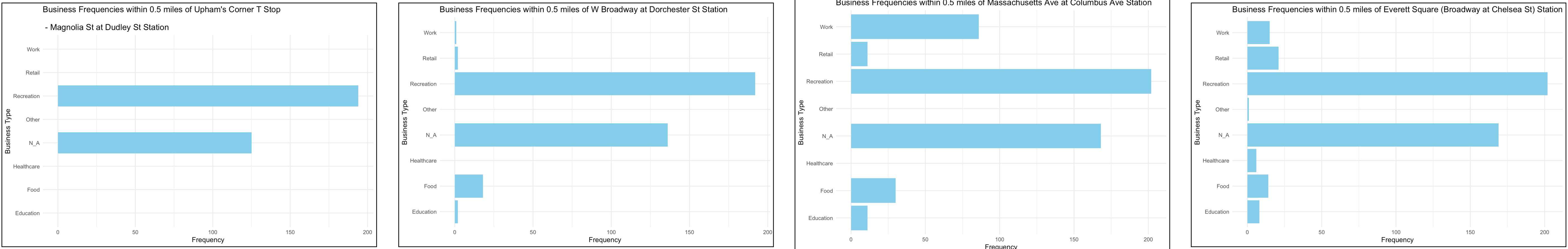
Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

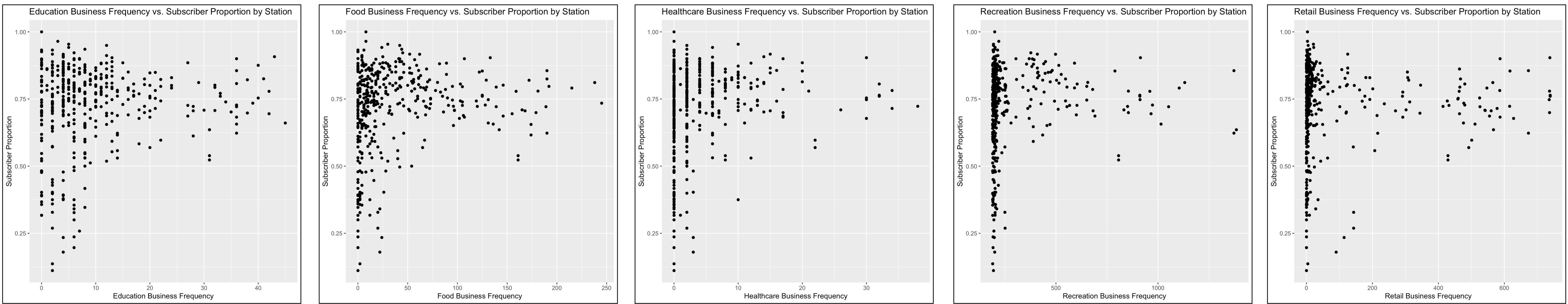
Methods

1. Download 2019 trip data from BLUEbikes
2. Use Python to get unique stations and calculate the subscriber proportion
3. Use R to query Overpass Turbo to get business data within 0.5 miles of a station
4. Use R to create scatterplots comparing business type frequency to subscriber proportion
5. Use R to run linear/multiple linear regression
 1. Use each business type separately and then combined
6. Compare model efficacy using R^2 and other indicators
7. Use R to run k-fold cross-validation to create a prediction model
8. Compare model efficacy using prediction error

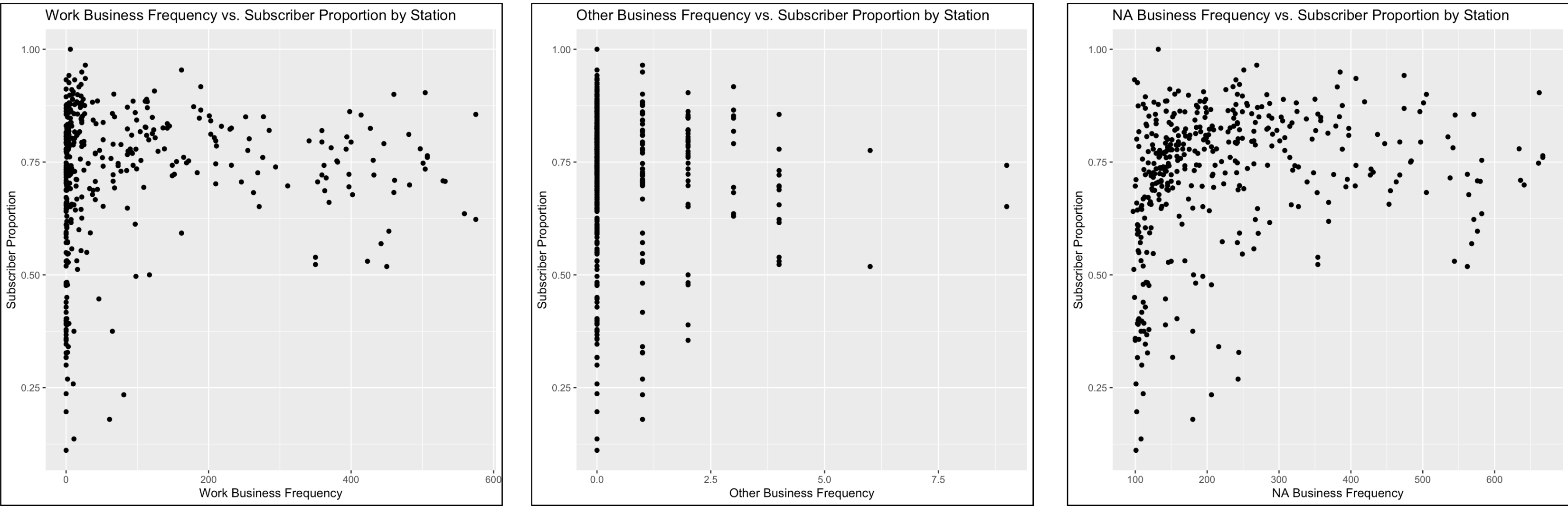
Results



Bar charts showing frequencies for 4 stations - Figure 1: Upham's Corner T Stop - Magnolia St at Dudley St Station; Figure 2: W Broadway at Dorchester St Station; Figure 3: Massachusetts Ave at Columbia Ave Station; Figure 4: Every Square (Broadway at Chelsea St) Station



Scatterplots showing business frequency with subscriber proportion



Scatterplots showing business frequency with subscriber proportion



Anova tables

Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Conclusions

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

References

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

